A Review of Current and Future Weather Data for Building Simulation

Manuel Herrera[a*], Sukumar Natarajan[a], David A. Coley[a], Tristan Kershaw[a], Alfonso P. Ramallo-González[b], Matthew Eames[c], Daniel Fosas[a] and Michael Wood[c]

[a] EDEn – Dept. of Architecture and Civil Engineering, University of Bath, Bath, UK
[b] Faculty of Computer Science, University of Murcia, Murcia, Spain
[c] College of Engineering, Mathematics and Physical Sciences, University of Exeter, Exeter, UK

[*]Corresponding author (Manuel Herrera) e-mail: amhf20@bath.ac.uk

Abstract

This article provides the first comprehensive assessment of methods for the creation of weather variables for use in building simulation. We undertake a critical analysis of the fundamental issues and limitations of each methodology and discusses new challenges, such as how to deal with uncertainty, the urban heat island, climate change and extreme events. Proposals for the next generation of weather files for building simulation are made based on this analysis. A seven-point list of requirements for weather files is introduced and the state-of-the-art compared to this via a mapping exercise. It is found that there are various issues with all current and suggested approaches, but the two areas most requiring attention are the production of weather files for the urban landscape and files specifically designed to test buildings against the criteria of morbidity, mortality and building services system failure.

Practical applications: Robust weather files are key to the design of sustainable, healthy and comfortable buildings. This article provides the first comprehensive assessment of their technical requirements to ensure buildings perform well in both current and future climates.

1. Introduction

Writing near the end of the first century B.C.E., the Roman architect Vitruvius, suggested that a building should seek to offer *firmitas*, *utilitas*, and *venustas* (firmness, commodity, and delight). *Utilitas* included the arrangement of spaces and the way the building provides both shelter from the external environment and comfortable internal conditions in which to carry out everyday tasks. Most buildings meet the basic requirements of shelter, so it is mainly the provision of comfort that taxes environmental designers. We spend 80-90% of our time inside buildings [1, 2] and poor internal conditions will not merely affect comfort but also impair occupant health and productivity. Since local weather and climate greatly affect the construction and performance of the building, good quality weather data is essential to simulate building performance. In this context, weather data should allow designers to stress test building performance for atypical conditions such as heat waves or cold snaps, since such conditions are more likely to cause performance failures. For example, the European heat wave of 2003 is said to have resulted in 70,000 excess summer deaths, primarily as a result of maladapted built environments [3, 4]. In the near future, the return period for such a heat wave is likely to change from 1 in 250 years to 1 in 50, or even 1 in 35 [5]. Since typical building lifetimes can be around 60 years or more, weather data needs to cover future changes, and be local to the building.

Weather files ideally need to:

1. Contain examples of typical conditions;
2. Contain examples of extreme conditions;
3. Be at the temporal resolution required by simulation packages (typically a 1 hour or higher resolution)
4. Be at a geographic resolution that matches changes in weather due to local topography in the country of interest;
5. Express the effect of the urban micro-climate; and
6. Contain examples of possible future climates, ideally considering the effects of climate change;

In addition to the six technical features above, weather files also need to be *credible*. This suggests a seventh feature necessary for the success of any weather file:

7. Proven track record with industry.

In this paper, we summarise the approaches and methodologies used to produce files that aim to match these requirements. This review is timely given that DEFRA and the UK Meteorological Office are about to expend a very large amount of computer time on producing the next set of climate predictions. However, the Directive on the Energy Performance of Buildings (EPBD) [6] has required built environment professionals to take measures that adapt planning policies and new building specifications to guarantee a minimum level of comfort and safety since 2002 [7]. This is especially important when considering the issues of providing adequate ventilation and in particular of limiting overheating [8] as this is associated with reduced worker productivity, morbidity and even mortality in vulnerable groups such as infants and the elderly [9, 10]. There is not a directive equivalent to EPBD in other world areas. However, the International Energy Agency (IEA) fosters the implementation of building energy saving policies at a national level in its member countries, but also in non-member countries such as China, India, and Russia.

This work is divided into the following sections. Section 2 reviews the state-of-the-art for creating weather data for building simulation. The different methodologies are classified according to how well the resultant weather meets the seven-point list of requirements. Section 3 focuses upon the methodologies for creating weather data from synthetic time series from a weather generator. Section 4 considers the creation of future weather data for building simulation both from synthetic data sets and from the morphing of historical observations. Finally, a discussion section considers the challenges that building engineers face when considering the future performance of buildings and increasing the resilience of the built environment to climate change.

2. Weather files for building simulation

The building industry mainly uses weather data to assess the design and performance of the built environment at the planning stage. This is becoming more important because climate change is likely to lead to an increase in the frequency of extreme weather events [11].

Dynamic building energy simulations were developed as early as the 1950s [12], but it was not until the energy crisis of the 1970s that the scientific community started using them to help improve the energy performance of buildings [13]. In early models, weather data were applied to the software packages in a variety of formats, but were standardised into ‘weather files’ by the 3rd generation of dynamic building simulators (following the classification of Clarke [14]). These generally take the form of typical weather years, created from hourly historic observations at a specific location [15]. However, the need to adapt buildings to the impacts of likely future climate change has created a requirement to incorporate climate change projections into these weather files, either by morphing the weather data or synthetically generating it [16, 17]. Weather files representing ‘extreme’ years (i.e. the selection of observed weather occurrences far from the norm) have also been introduced to analyse a building design’s response in case of severe weather conditions [18].

2.1. Format of weather files

According to the bi-annual publications of the International Building Performance Simulation Association (IBPSA), the most commonly used building simulator is EnergyPlus. As well as being a stand-along programme, EnergyPlus is also the basis for more sophisticated software (e.g. DesignBuilder), which are capable of analysing not only energy consumption, but water use and daylighting.

The popularity of EnergyPlus has driven the popularisation of its native weather file format - the ‘.epw’ file (**E**nergy **P**lus **W**eather). These .epw files are text-based CSV files that contain a years-worth of hourly weather variables for a given location. The file structure was developed by USDoE, and is also used in other software such as ESP-r, IES, and TAS. The USDoE offers 2,590 weather files for different worldwide locations in the current climate.

Figure 1 shows the header of a typical .epw weather file, the first line of the header shows the location and the specific format of the weather data (in this case TMY2). The header also contains the location, longitude, latitude, time zone, elevation, annual design conditions, monthly average ground temperatures, typical/extreme periods and holidays/saving periods. It also has information on which data periods were included.

```
1  LOCATION,BOULDER,CO,USA,TMY2-94018,724699,40.02,-105.25,-7.0,1634.0
2  DESIGN CONDITIONS,0
3  TYPICAL/EXTREME PERIODS,6,Summer - Week Nearest Max Temperature For Period,Extreme,7/13,7/19,Summer
   Period,Typical,8/17,8/23,Winter - Week Nearest Min Temperature For Period,Extreme,12/ 8,12/14,Winte
   2,Autumn - Week Nearest Average Temperature For Period,Typical,10/27,11/ 2,Spring - Week Nearest Av
4  GROUND TEMPERATURES,3,.5,,,,-0.30,-1.30,0.47,3.24,10.48,16.07,19.85,20.99,19.04,14.70,8.90,3.51,2,
   4,,,,6.41,4.58,4.19,4.74,7.48,10.43,13.14,15.02,15.47,14.40,12.05,9.18
5  HOLIDAYS/DAYLIGHT SAVINGS,No,0,0,0
6  COMMENTS 1,TMY2-94018 -- WMO#724699
7  COMMENTS 2, -- Ground temps produced with a standard soil diffusivity of 2.3225760E-03 {m**2/day}
8  DATA PERIODS,1,1,Data,Sunday, 1/ 1,12/31
9  1970,1,1,1,60,B8E7B8B8*0?0?0?0?0?0?0?0B8B8B8B8*0*0F8F8A7E7,-7.0,-8.5,88,83400,0,0,238,0,0,0,0,0,0,
10 1970,1,1,2,60,A7A7A7A7*0?0?0?0?0?0?0?0A7A7A7A7A7A7F8F8A7E7,-7.2,-8.3,91,83400,0,0,238,0,0,0,0,0,0,
11 1970,1,1,3,60,B8E7B8B8*0?0?0?0?0?0?0?0B8B8B8B8*0*0F8F8A7E7,-8.1,-8.8,94,83400,0,0,228,0,0,0,0,0,0,
12 1970,1,1,4,60,B8E7B8B8*0?0?0?0?0?0?0?0B8B8B8B8*0*0F8F8A7E7,-9.1,-9.4,97,83400,0,0,220,0,0,0,0,0,0,
13 1970,1,1,5,60,A7A7A7A7*0?0?0?0?0?0?0?0A7A7A7A7A7A7F8F8A7E7,-10.0,-10.0,100,83300,0,0,213,0,0,0,0,0
14 1970,1,1,6,60,B8E7B8B8*0?0?0?0?0?0?0?0B8B8B8B8*0*0F8F8A7E7,-10.0,-10.1,99,83400,0,0,213,0,0,0,0,0,0,
15 1970,1,1,7,60,B8E7B8B8*0?0?0?0?0?0?0?0B8B8B8B8*0*0F8F8A7E7,-10.0,-10.3,97,83500,0,0,213,0,0,0,0,0,0,
16 1970,1,1,8,60,A7A7A7A7*0E5E4E5I5I4I5I6B8A7A7A7A7A7F8F8A7E7,-10.0,-10.6,96,83600,71,755,213,24,69,1
```

Fig. 1. Sample of .epw weather file, showing header information and weather data form line 9.

2.2. Files for typical weather conditions

Files for typical weather conditions include hourly data on temperature, dew point, global horizontal radiation, diffuse solar radiation, wind speed and wind direction. These files are used to estimate the average building energy use and carbon emissions [19, 20]. A typical weather file is created from historic data (usually around 20-30 years of data, depending on the data availability). This data is compiled by comparing the cumulative and the empirical distribution functions of different meteorological variables within the base data set. The number and weighting of different meteorological variables considered is a feature of the weather file type (i.e. TMY, TRY, etc. – see below). Table 1 shows a representative sample of typical weather files used around the world. It is worth mentioning that despite having different sources or ways in which weather files are created to form distinct file types, several of them use common file formats such as the EPW format.

Table 1. A short list of weather file types in various countries. The period depends on the data availability at the location[1].

Acronym	Complete name	Region	Sites	Period
RMY	Representative Meteorological Year	Australia	69 locations	1967-04
CSWD	Chinese Standard Weather Data	China	270 locations	1982-97
ISHRAE	Indian Typical Years from ISHRAE	India	62 locations	1991-05
IGDG	Italian ‘Gianni De Giorgio’	Italy	68 locations	1951-70
SWEC	Spanish Weather for Energy Calculations	Spain	52 locations	1961-90
UK TRY	Test Reference Year (CIBSE)	UK	14 locations	1984-13
TMY	Typical Meteorological Year	USA and others	1,020 locations	1991-05
WYEC	Weather Year for Energy Calculations	USA/Canada	77 locations	1953-01
IWEC	International Weather for Energy Calculations	Worldwide	3,012 locations	1991-05

There are two ways to construct a typical weather year. The first is by identifying a continuous 12-month period as typical. The second is by applying the ranking criteria to individual months from the basis set, which are then assembled into a composite 12-month year. The UK TRY and TMY both use the latter approach and are computed using the Finkelstein-Schafer (FS) statistic [21]. This means that each month in the file might be from a different year. Comparisons of these composite years with the basis set indicate that both the UK TRY [20] and TMY (with the updated file formats to TMY2 and TMY3) [22] have advantages over a single year approach.

In the following a summary of the characteristics of the most representative composite weather year files (TRY, TMY, and IWEC) and some of their extensions and updates is given.

- The Test Reference Year (TRY) was developed in 1976 for 60 locations in the United States [23]. The baseline period was from 1948-1975. From this baseline, years with monthly extreme values were filtered out until a single year remained containing the least severe (or most average) weather conditions. TRY initially contained dry bulb, wet bulb, and dew point temperatures, wind direction and speed, barometric pressure, relative humidity, cloud cover and type. Later on, the TRY methodology was modified [24, 25] and its scope was expanded to generate a complete weather data set for several locations worldwide. The exact modification of the the TRY depended on the institution that created the files. The differences included the weighting of relevant parameters and even the inclusion or not of one or more parameters. One example is the Danish Design Reference Year (D-DRY) [26], which includes specific parameters, such as 5-minute values for direct normal radiation, or forecast information to be used for the simulation of energy management systems. To create a D-DRY the data set of basis months are ranked according to the distance (measured in standard deviations) of each variable per month from the value of the long-term mean.

  The Chartered Institution of Building Services Engineers (CIBSE) in association with the UK Met Office produces CIBSE TRYs for the United Kingdom [27]. In the case of the UK

1 The information of various weather files is regarding to their corresponding updated formats. This is the case of TMY, updated to TMY2 and TMY3; WYEC, updated to WYEC2; and IWEC, updated to IWEC2.

TRY originally each of the 3 environmental parameters carries an equal weighting namely dry bulb temperature, cloud cover (used as a proxy for solar irradiation), and wind speed, this was deemed the most appropriate for the naturally ventilated buildings typical of the UK [18]. In the recent update [27] the environmental parameters have been updated to use dry bulb temperature, cloud cover (used as a proxy for solar irradiation), and relative humidity as primary variables with wind speed as a secondary variable. A comparison of various TRY approaches is in Bilbao et al. [28] while an analysis of UK TRY is in Eames et al. [27]. The UK TRY is composed of 12 separate months of data each one chosen to be the most average month among a set of years. The cumulative distribution functions on which the UK TRY is based are made up of the daily mean of values of dry bulb temperature, cloud coverand relative humidity. These daily means are computed using hourly values from all the months of the basis years considered. Component months are chosen using the FS statistic method, essentially, the three months with the most average values of temperature, radiation and relative humidity are selected. From these three months, the month with the most average wind speed is then selected for the UK TRY.

- The Typical Meteorological Year (TMY) [29] is also based on the FS statistic method over the data sets derived from the years 1961-1990. However, the TMY uses more input variables than the TRY: minimum, maximum, and mean values of dry bulb temperature, dew point temperature; and minimum and maximum of wind speed. TMY, in addition to global radiation, also includes direct normal radiation. Details of the concept and work arising from the TMY can be found in Lam et al. [30], Ecevit et al. [31] and Yang et al. [32]. The use of Typical Meteorological Years [33] is common in the U.S., but is also being considered by other countries.
    - The Typical Meteorological Year 2 (TMY2) [34] and Weather Year for Energy Calculations 2 (WYEC2) [35] are similar to TMY but with more complex solar models. In addition, weightings for dry bulb and dew point temperature are changed slightly to give more emphasis to dry bulb and dew point temperatures and less to wind speed. The base time period for TMY2 is 1961-1990.
    - The Typical Meteorological Year 3 (TMY3) [29] is created using a similar procedure to TMY2. The TMY3 is based on 15 basis years (1991-2005), but at sites where data is available for 30 years, the base time period spans 1976-2005. The Typical Principal Component Year (TPCY) [32] is an alternative to the TMY designed to reduce the data set of the TMY to a set containing a small number of uncorrelated components. The new database is formed by artificial variables generated using principal component analysis (PCA). Each component is a combination of the original variables representing a large fraction of the variability in the original data; that is, reducing the database dimensionality while keeping a maximum amount of information.
- The International Weather for Energy Calculations (IWEC) year [36] is an attempt, by the American Society of Heating, Refrigerating and Air-Conditioning Engineers (ASHRAE), to internationally unify weather files. Given the huge availability of raw weather data on the Web, the strategy behind IWEC is to search the entire Integrated Surface Hourly (ISH) weather data; producing as many “typical year” weather files as possible. Then to release these in a similar format to TMY3. IWEC files contain weather observations of wind speed and direction, sky cover, visibility, ceiling height, dry-bulb temperature, dew-point temperature, atmospheric pressure, liquid precipitation, and current weather for at least 12

years of records but up to 25 years. If the ISH database does not have measurements of solar radiation, then the hourly global horizontal radiation and the direct normal radiation is estimated empirically using the other variables such as sun-earth geometry, cloud cover, relative humidity, temperature difference from three hours previously, and wind speed. IWEC has now been updated to IWEC2 [37] to contain slightly more cooling degree-days. There is also substantially more variation in the solar radiation. The IWEC2 weather files have lower weights for global horizontal radiation but higher weights for direct normal solar radiation than the previous version. There are currently 3,012 IWEC2 files for locations outside the USA and Canada; TMY3 files are available within USA and Canada [38].

2.3. Files for extreme weather conditions

The typical weather years introduced above are based upon identifying periods of average weather over the basis years. As such these methods are suitable for the calculation of model outputs such as typical annual energy use. However, using average data to produce weather files provides no information about the natural variability of the weather, which is of primary importance when trying to consider extreme or atypical weather conditions. This is needed because extreme weather events such as, heat waves, drought periods, or cold snaps are essential for modelling overheating in buildings [39], analysis of thermal comfort [40] or estimating peak energy use [41]. Extreme weather events are likely to become more frequent in the future as a result of climate change [11, 42], hence the building engineering community is attempting to adapt building design and occupant behaviour to cope with such extreme events. It is possible to transform typical weather years using a variety of methods. Pernigotto et al. [43] weight the weather parameters involved in TRY computation in different ways depending on if the weather file is intended for heating or cooling analysis following the European technical standard, EN ISO 15927-4. A similar approach is adopted by Kalamees et al. [44], who use weighting factors depending on climate zone variables to modify EN ISO 15927-4 in the computation of heating and cooling energy demand in buildings.

As an alternative to these approaches, there have been specifically designed years to reflect less typical years, for example warmer than average summers. In the following the most important are described.

- The Design Summer Year (DSY) [45] is primarily an attempt to estimate the impact of warmer than average summers and was initially intended primarily for the sizing of mechanical cooling systems. The DSY is the year that falls in the middle of the upper-quartile of the basis years' dataset, ranked according to summertime (April to September) average dry bulb temperature, this is generally the third warmest summer for a basis dataset of 20-years. The DSY does not take into account extreme temperatures in individual months or incident solar radiation, both of which are of great significance for assessing the overheating performance of buildings [39]. This means that periods of high temperature (such as heat waves) in relatively cool summers are not considered. This is a problem, as summers such as 2003, which resulted in so many deaths across Europe are often not ranked highly in the basis set when considering average summertime temperature. There are a number of variations of DSY attempting to improve its sensitivity to severe weather conditions, discussed below:
    - The Probabilistic Design Summer Year (pDSY) [46, 47], is based upon different overheating metric to the DSY, namely by selecting years on the basis that they contain events warm enough to cause a degree of overheating within in a *notional* building that maintains almost the same temperature as the external temperature. In

TM49 years are ranked according to a new metric called ‘weighted cooling degree-hours’ (WCDH), which is the difference between the operative temperature[2] and the adaptive comfort temperature [48, 49]. Since the greatest values of WCDH are likely to be observed for large return periods, an extended set of basis years are used (1950 to 2006). For the rest of the UK, a further new metric named the Static Weighted Cooling Degree Hours was used with the reference temperature equal to the 93rd centile of dry bulb temperature – the temperature at which the deaths due to overheating can be attributed. However, given the low frequency of occurrence for extreme events even over this range of years, it has been proposed that a synthetic weather generator [50] might be a better option as this could provide a very large number of basis years.

  - The Summer Reference Year (SRY) [51] is based upon the TRY methodology and is intended to represent near-extreme conditions. The SRY adjusts the TRY dry bulb temperatures to values representing the 90th percentile of a temperature distribution function created from the summer months of the basis dataset. Wet bulb temperature, wind speed and atmospheric pressure are consequently changed after this initial adjustment. Solar irradiation follows a similar two-stage process as used for producing the dry bulb temperature. Ultimately, the SRYs should contain at least one warm spell period to be useful in overheating computations.
  - The near extreme Design Reference Year (DRY) [52-54] uses a three-fold process to create a weather year. The process requires a large amount of weather data; hence the UKCP09 weather generator is used to produce 3000 basis years. The individual months are ranked according to monthly mean temperature. Twenty years’ worth of these ranked months (20 Januarys, Februarys, etc.) each centred on the middle of the upper quartile (87.5%) are identified and using the FS statistical method the three months with the lowest combined rank sum of dry bulb temperature, humidity and irradiance are selected. The month within this group of three with the closest mean monthly wind speed to the 20-year average is then chosen. This process is repeated for all 12 months and the DRY is created. This method has the benefit that the process can be used for current or future weather and the method can be altered to favour humidity or irradiance instead of temperature.

- The Extreme Meteorological Year (XMY) [55] is an extension of the TMY idea over the whole year and uses the same weather variables and similar weightings as the TMY. The XMY is based upon choosing extremes from the basis set instead of averages. Months with the highest and lowest hourly averages values in the basis years (1999-2013) are combined to form a year with the hottest summer and the coldest winter.
- The Untypical Meteorological Year (UMY) [56] is based on the WYEC2 with altered weight parameters to compose the meteorological year. In the UMY the parameters related to maximum and minimum dry bulb temperatures, solar radiation, and wind speed were proposed as the most important. Other weight combinations are possible and its use is suitable depending on the climate of the region in which the weather file is computed [56]. UMY files produce comparable results to TMY2 in regular conditions (where a regular

2 The operative temperature is defined as the uniform temperature of a radiantly black enclosure in which an occupant would exchange the same amount of heat by radiation and convection as in the actual non-uniform environment. For typical indoor conditions with low air speeds, this equates to the arithmetic mean of the air and mean radiant temperatures.

weather file would be sufficient) but improves the prediction of the maximum energy use during severe weather events.

- The Hot Summer Year (HSY) [57] has two versions: HSY-1 based on the highest WCDH year when checking the summers (June, July and August) from 1975 until 2006, and HSY-2 based on the year of summer with the most hours of Physiologically Equivalent Temperature (PET) [58] over 23 °C, using the same year basis: 1975-2006.

2.4. Limitations of using observed data for typical and extreme weather files

Observed (i.e. historical) weather data are the principal source of data for all the weather files, listed in Sections 2.2 and 2.3 above. There are two principal sets of limitations in the use of such data. The first applies to both typical and extreme weather files, and can be summarized as follows:

- Weather files are based on a relatively small number of weather stations with a heterogeneous spatial distribution. The applicability of this data far from the observation site is reduced. Furthermore, as a consequence of the scarcity and irregular spatial distribution, adjacent locations might frequently use different sets of observed years for the generation of weather files. For example, moving a potential building 100m on a site might mean crossing a weather file boundary, with significant changes in predicted performance.
- As the weather files cover large areas, it is also possible for weather files for coastal locations to be applied to inland and upland sites or vice-versa, potentially resulting in incorrect and expensive design decisions to ensure compliance with building regulations or guidance [59].
- Airports are a common location for weather data collection and there may be clear differences between the environment around such locations and other areas, such as nearby cities.
- A consequence of the use of historic weather is that, by definition, the modeller is looking at how the building would have performed given a times series that will never be repeated, not how it will perform in any year that the building might experience after it is constructed. Furthermore, the process used to form standard weather files of stitching the most typical January to the most Typical February, etc., produces a very unlikely time series as these months very rarely occurred successively. However, it is likely that a typical weather file processed from a recent twenty to thirty year time series will be representative of the next decade. At least for annual energy use in a heating dominated climate such as the UK, this approach does give reasonable results in comparison to modelling the building using the complete set of basis years and then picking the average result [60].

The second set of limitations is specifically related to weather files representing extreme weather conditions. These can be summarised as follows:

- The primary limitation of extreme weather files is how to overcome the fact that extreme weather is, by definition, a low frequency event. So, a time series collected to create the typical weather years are not long enough to provide sufficient information about the frequency and intensity of extreme weather events that might only occur once every 50 or 100 years. Also characteristics of weather time series change over time, further complicating the extreme events analysis [61].
- The use of historical data implies that no information is contained on how climate change may affect weather patterns and extreme weather events.

- There is also cause for concern about the variables chosen to compile the weather files: for example, the DSY does not guarantee correlation with the building internal environment because this also depends on other variables (wind speed, cloud cover, etc.) which are not considered when creating this weather file [39]. The pDSY approach represents a fair advance beyond the simple DSY, but again this weather file does not consider important variables that determine the performance of buildings in hotter summers such as wind and solar radiation. The DRY is more comprehensive, however, there are still issues regarding replicating the natural variability and weather extremes using the weather generators on which this is based. In addition. As an alternative the SRY does consider solar radiation as an independent parameter, however, it excludes wind.

3. Synthetic weather

In this section, we introduce weather generators, list the most widely used types of generators, and conclude with some limitations of synthetic weather. Given the limitations of historical weather data for the construction of weather files for simulation, interest has grown in the use of synthetic weather data that is able to mimic weather behaviour. Synthetic weather data can provide a valuable aid in the formulation of policies and in the decision-making process by reinforcing the information available at coarser spatiotemporal scales and providing insight in the case where no data is available. Synthetic weather can also simulate extreme conditions which while being statistically representative of the location, have not been observed.

Weather generators use computer algorithms that produce a long time-series of weather variables with statistical properties comparable to existing historical records. Weather generators can also simulate meteorological variables at different time scales, on the basis of empirical statistical models, based upon the downscaling of an ensemble of climate model outputs [62-64]. Weather generators are often developed in two steps: firstly, by modelling of daily precipitation (see Figure 2) and then generating the remaining variables of interest based upon the rain occurrence. These variables often are maximum and minimum temperature (T_{max}, T_{min}), and solar radiation (R). Other variables such as wind direction and speed are then derived from the key variables. The decision of either a wet or dry day depends on the amount of rainfall observed; dry is used if the rainfall is below a certain threshold (usually, 0.02 mm); otherwise the day is classified as wet. For each month different model parameters are used in order to represent the seasonal variations in both the magnitude of climate variables and their cross-correlations (correlations between single variables over different periods of time).

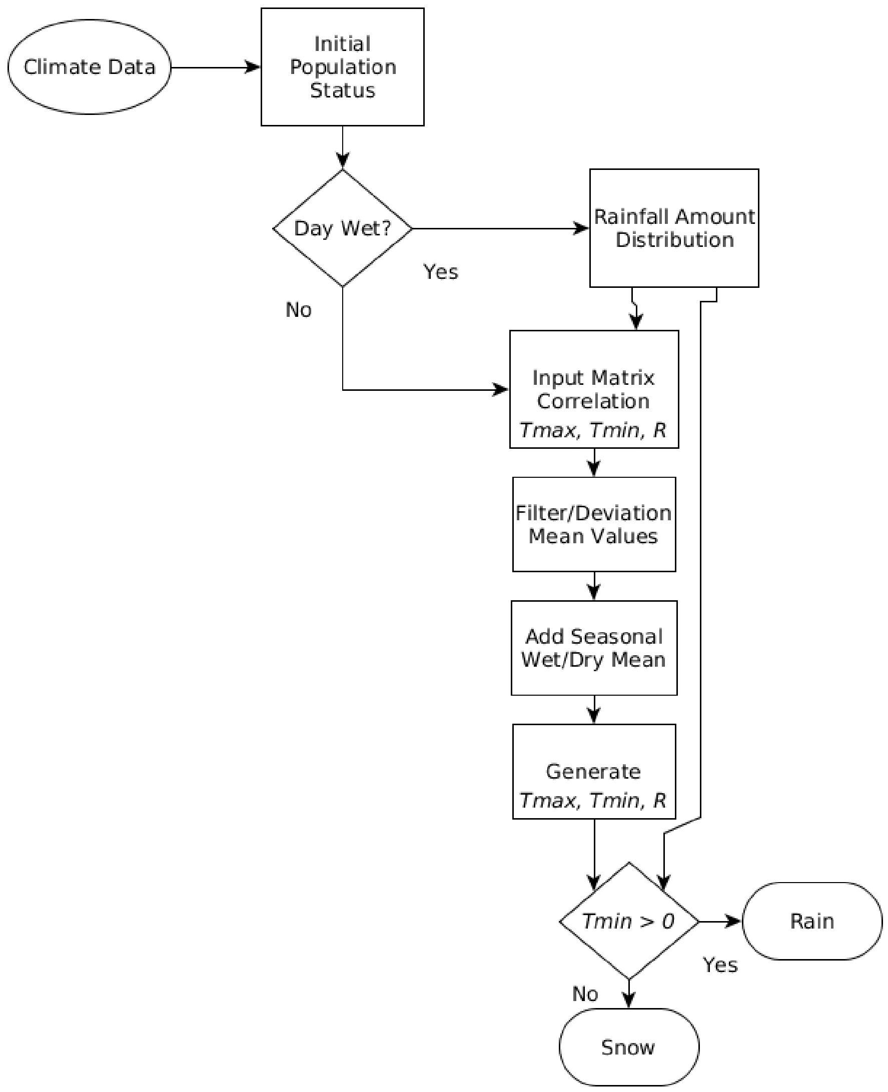

Fig. 2. Flowchart of different stages for modelling daily precipitation in a typical weather generator

3.1. Most widely used weather generators

Weather generators can be classified depending on the assumptions made of the variables on which they are based. The majority of the most widely used weather generators are parametric. That is, they involve assumptions about the statistical properties and distribution functions of the inputs. Nonparametric weather generators are data-driven. The nonparametric weather generators often use resampling and simulation methods that do not need to meet any inherent data assumption. A weather generator which combines the two options is called a semi-parametric weather generator.

A list of the most common parametric weather generator models are described below:

- WGEN [65] describes the serial dependence of precipitation, p, using a sequence of possible events in which the probability of each event depends only on the state attained in the previous one (Markov chain). The T_{max}, T_{min}, and R values are related to p based upon the wet or dry status of the day. The dependence among these three variables is preserved using cross-correlation coefficients of these variables at different points in time. The WXGEN model [66] is introduced as an adaptation of WGEN for use in soil erosion impact calculations.

- The CLIGEN model [67] advances WGEN by producing individual storm parameter estimates, including time to peak, peak intensity and storm duration. This feature allows CLIGEN to be used for soil erosion models [68, 69]. The model reproduces historically accurate monthly distributions of its parameters. These main inputs are: rainfall, temperature, and wind information, along with a number of soil and topographic inputs. However, it is not appropriate for generating daily weather, due to the generation of key variables not being cross-correlated with precipitation. As a result, variables involving wet-days and solar radiation are not reproduced.
- ClimGen [70] uses a similar approach to WGEN. However, the daily precipitation amounts are assumed to follow a Weibull distribution as opposed to the Gamma distribution as used by WGEN. ClimGen generates half hour precipitation intensities using the assumption that precipitation amounts within a storm are exponentially distributed. This makes ClimGen more suitable for estimating extreme rainfall events [71, 72].
- Met&Roll [73] uses the standard deviations of extreme temperatures (T_{max} and T_{min}) and daily sums of global solar radiation (R) to produce synthetic weather time series for hydrological modelling and crop production assessments. The process is linearly dependent on previous values and on a stochastic term based on first-order autoregressive models (AR1) with an annual cycle and monthly seasonality.
- CRU-WG [74, 75] is an improvement upon the WGEN structure. CRU-WG uses the basis years 1961-1990 to estimate the parameters used to calibrate the stochastic model, which generates the daily weather variables. The first variable generated is precipitation, other variables such as T_{max} and T_{min}, pressure, wind speed, and sunshine hours, are generated from their observed cross-correlations with precipitation. Other CRU-WG results; such as relative humidity and potential evapotranspiration, among others are derived from these variables. CRU-WG better represents low probability events such as high temperatures and extreme precipitation.
- WeaGETS [76, 77] is a Matlab-based stochastic daily weather generator for producing daily precipitation, maximum and minimum temperatures series of unlimited length, thus permitting impact studies of low frequency meteorological variables. WeaGETS has the advantage of incorporating the computational schemes of other well-known weather generators, as well as offering unique options, such as correction of the underestimation of inter-annual variability and the ability to use Markov chains of varying orders. More importantly, the use of Matlab allows for easy modification of the source code to suit the specific needs of users. It would be straightforward, for example, to add different precipitation distribution functions.
- WETTREG [78] selects blocks of data from the observed time series based upon classifying circulation patterns that have a strong link to the local regional climate. The average daily temperature is calculated using a 5-day running average. For example, the temperature on the 22nd December replaced by the average temperature of December 20th to 24th. The time series is analysed looking for episodes (e.g. temperature above certain threshold) to identify weather patterns within the climate model data. This is used for the stochastic generation of the other variables. The data is then interpolated to create a complete regionalised time series.

- AnaWEGE [79] is a stochastic weather generator based on an ‘analogues of circulation’[3] downscaling methodology that compares atmospheric states, from General Circulation Models (GCM) data [80]. The similarities produced by these ‘analogues of circulation’ specify a coherent physical and spatial relationship between a set of locations in which similar large-scale patterns provide suitable initial conditions for generating weather variables [81].
- SUWG, spatialized urban weather generator [82], is designed to simulate the urban heat island effect. SUWG extrapolates data collected at the weather station at a 30m height over the city, to study the weather characteristic of neighbourhoods. Then, a 2D grid mesh is applied over the whole city under consideration and heat flux effects of the urbanisation in each of the cells are estimated. By taking into account these urban area effects on the meteorological parameters, it is possible to use the model to obtain data at a suitable height (e.g. 2 m) and adapt SUWG for building energy analysis.
- Meteonorm [83] extrapolates hourly data from statistical data for a location. Where statistical data are not available, Meteonorm interpolates from other nearby sites. Meteonorm is a combination of a climate database, a spatial interpolation tool and a stochastic weather generator, with global radiation data obtained from the Global Energy Balance Archive (GEBA). This allows typical years with hourly or minutely time resolution to be created for any site [84]. The stochastic generation of global radiation is based on a Markov chain model for daily values and an autoregressive model for hourly or minutely data. Meteonorm has been extensively used for generating typical meteorological years [85].

The most common semi-parametric weather generator model is described as follows:

- LARS-WG [86] is a semi-parametric version of WGEN, which uses a mathematical representation used for a daily weather simulation of the process where the model parameters are directly estimated from the sample data. This has the advantage of requiring no prior assumptions of probability distributions of the parameters. LARS-WG uses semi-empirical distributions for the duration of wet and dry day series, daily precipitation and daily solar radiation; this means LARS-WG more closely matches observed data than WGEN. However, both weather generators fail when considering variation in the correlation between variables. This is of importance mainly in the modelling accuracy for estimating weather scenarios where two or more variables interact with each other.

The most common non-parametric weather generator models is described as follows:

- KnnCAD version 4 [87] is a non-parametric weather generator algorithm for precipitation and temperature based on spatial rainfall simulations created through associated K-nearest neighbours’ weighting. Given that KnnCAD uses non-parametric models for its inputs, it is not necessary to make further assumptions regarding the spatial correlations and the probability distributions for each variable of interest. Consequently, KnnCAD may be applied within a more general framework than parametric and semi-parametric weather generators. This allows the use of resampling techniques to obtain an appropriate temporal correlation for the analysis of low frequency events.

---

3 Analogues of circulation are periods of time which have similar atmospheric circulations to the days of interest. These can be obtained in different ways, depending on the target variable, its domain and the reference period used to compute them.

3.2. Limitations of weather generators

The primary limitation of weather generators is that they are based upon statistics derived from historical observations of weather. This means that it is unlikely that extreme events will be correctly represented, as these will rarely, if at all, have occurred during the short historic record used to create the generator. In addition, there is the inherent assumption that future weather patterns will be the same as those observed historically.

The output from weather generators is synthetic data, thus it is necessary to investigate the output to assess their representativeness. There are several evaluations of weather generators in the literature: Soltani et al. [88] assessed WGEN capabilities. WGEN was found to be representative of both current and future weather. Semenov et al. [89] carried out a comparison between WGEN and LARS-WG for diverse climates. Both generators were found to not reproduce the variability of the data and also underpredicted both heat waves or cold snaps. Hayhoe et al. [68] analysed CLIGEN for the use with the Canadian climate. CLIGEN provides good results with observed mean values but underestimate their variation.

In theory, weather generators can be used to generate enough data to evaluate the probability of extreme weather events [90]. Observed time series represent one single ‘realisation’ of the climate, whereas a weather generator can create many ‘realisations’ and hence, potentially, a wider range of feasible situations. Smith and Hanby [91] developed methodologies for generating DSYs based upon iterative weather generator simulations coupled with multiple sampling (up to 100 times) the 30-year period used as the baseline for the predictions. However, there is still a problem with regards to embedding heat waves and cold spells into weather generators, which are generally primed with average data and often constrained by the use of precipitation as input.

There are several studies in the literature examining extreme weather events [50, 63, 92, 93] which might help the applicability of generation of weather files for building simulation. Chen and Brisette [94] compared 5 weather generators (WGEN, ClimGen, CLIGEN, WeaGETS and LARS-WG) for a dry climate case-study at 54 stations at the Loess Plateau (China). These comparisons were based on reproducing precipitation, minimum and maximum temperatures for various scenarios. As precipitation does not follow a Normal distribution, the comparisons between the weather generators and the observed data were approached by nonparametric statistics.

LARS-WG was found to be better at simulating the overall distribution of daily precipitation, especially at preserving the the skewness and kutosis but was worse than markov chain models for reproducing extreme wet events especially where the observations of such extremes were limited. CLIGEN and WeaGETs performed betted than WGEN and ClimGen at simulating daily precipitation amounts with WeaGETs the best at simulating extreme precipitation events. Regarding temperatures, the t-tests showed that neither the observed maximum nor minimum temperatures are significantly different from the generated data at level 0.05 for all 54 stations for all weather generators, apart from WGEN, for 4 stations simulating minimum temperatures**.** However, the standard deviation was not reproduced for most of the weather generators; the F-test showed that only CLIGEN was statistically similar to observations. Overall WeaGETS had the best performance with respect to reproducing both precipitation and temperatures over the case study region.

4. Future weather

Climate change is expected to result in an increased frequency and intensity of extreme weather events. It would directly affect key areas for the Nature and human being such as agriculture [95], water resources [96] and energy systems [97]. The built environment will also go through climate change threats such as those related to overheating issues.

The Intergovernmental Panel for Climate Change (IPCC) has studied climatic change using a number of possible future emissions scenarios. The different scenarios were created by varying the rate of greenhouse gas production using different socio-economic scenarios [98, 99]. As a result of the IPCC fifth Assessment Report (AR5) [100], these emissions scenarios have evolved into *representative concentration pathways* (RCPs). These RCPs now provide the input to both Global Climate Models (GCMs) and finer resolution Regional Climate Models (RCMs) [101-103].

The results obtained from the GCMs and RCMs represent averages over regions or numerical grids, with the size of these grids depending on the model resolution. Even for the highest spatial resolution GCMs, the numerical grid is still too coarse to allow a suitable understanding of the impacts of climate change on the built environment. The finer scale and implied increase in geographic and land surface information of RCMs allows for greater geographic resolution. This enables RCMs to provide weather and climate data at horizontal resolutions up to 50 or even 25 km. Projects such as ENSEMBLES [104], NARCCAP [105] and CORDEX [106] used combinations of several RCMs to provide higher resolution final models (achieving 10-20 km of horizontal resolution). An alternative to RCMs is to use statistical techniques to ‘downscale’ GCM models to a finer grid. These techniques directly incorporate observations collected at weather stations and those collected through one or multiple GCMs (by combining various scenarios). This technique assumes that large-scale meteorology and geographic features influence local weather and climate. The climate model data may also need to be temporally downscaled (time scale adjustment) if the resolution is too coarse for the required purpose of the final application.

Both RCM and GCM downscaled data can either be incorporated into a weather generator, which will produce future synthetic series of weather data [86], or by means of a mathematical transformation (morphing) produce a future time series based upon historic weather observations [107]. The creation of future weather files can be approached in two ways [17]: by combining climate projections with a weather generator to allow the creation of typical future weather years [108], or by a mathematical transformation (morphing) of the time series of existing current weather files using climate change anomalies from a GCM or RCM [109].

4.1. Climate projections

Each climate model projection is the result of many iterative simulations of climate models (termed a multi-model ensemble). These ensembles are combined to produce the projections of future climate. The following items list a sample of widely used international climate models, such as the Coupled Model Inter-comparison Project (CMIP), and a number of climate models specifically developed in the UK:

- CMIP3/5. Coupled Model Inter-comparison Project Phase 3 and Phase 5 [110, 111]. Under the World Climate Research Programme, the Working Group on Coupled Modelling established the Coupled Model Inter-comparison Project (CMIP) as a standard experimental protocol for studying the output of coupling comprehensive three-dimensional atmospheric GCMs within oceanic general circulation models, with sea-ice models and models of land-surface processes. Both CMIP3 and CMIP5 have been used to generate projections of future global climate conditions using a large number of GCMs. CMIP3 and CMIP5 cannot be directly compared due to the different approaches used for estimating future greenhouse gas emissions. CMIP5 uses the Representative Concentration Pathways of greenhouse gases [112], whereas the older CMIP3 is based on the SRES [101] emission scenarios as used in IPCC, Fourth Assessment Report, AR4 [110]. There are several approaches to downscale CMIP3 and CMIP5 models, with most achieving a spatial resolution of approximately

50km, depending on the ensemble used for running each model [113-115]. ENSEMBLES [104] and EURO-CORDEX [106] are the European alternative based on an intensive use of RCMs. The next generation of CMIP: Phase 6 (CMIP6) focuses on identifying the origins of model biases and the assessment of future climate change given climate variability, predictability and the uncertainties in scenarios.

- UKCIP02/UKCP09. UKCIP02 (UK Climate Impacts Programme 2002) climate projections [116] were based on a series of climate modelling experiments completed by the Met Office's Hadley Centre (UK) following 4 of the emissions scenarios published in 1990 by IPCC for three future time-slices [99]. They comprise of climate projections at a spatial resolution of 50 km in the case of UKCIP02 and 25 km for UKCP09 (UK Climate Projections 2009). The approach relies on deriving factors of change for various statistics, rather than using RCM rainfall climatology directly [117, 118]. UKCP09 [119, 120] represents an evolution of the projections achieved by UKCIP02. It is based on a post-processing, phase-transforming simulation of results into probabilistic projections, allowing modelling of the uncertainties associated with the projections [121, 122]. The probabilistic projection methodology in UKCP09 involves sampling climate modelling uncertainties by combining results from perturbed variants of the HadCM3 configuration of the UK Met Office global climate model with projections from an ensemble of alternative international climate models. By comparison, UKCP09 follows only 3 of the IPCC emissions scenarios but for an increased number of time periods, with decadal time-slices from 2020-2080. There are a number of applications arising from the use of UKCIP02 and UKCP09, among them we highlight the project BETWIXT for building simulations and the weather generator EARWIG. The following points introduce these two interesting outcomes:
    - BETWIXT [123] is a project based on the use of UKCIP02 and CRU-WG to support the provision of daily and hourly meteorological data useful for building applications. A major advance in BETWIXT is fitting models to current and projected future rainfall statistics within the RainClim software package by using factors affecting rainfall derived from an RCM output. Then the hourly rainfall information is aggregated to a daily basis to firstly, generate daily temperature and then run hourly regressions for temperature and other variables. BETWIXT puts special emphasis on dealing with extremes and so third moment order models are used to better fit the model [124].
    - EARWIG [118] is a weather generator originally used with UKCIP02 and then iterated for use with UKCP09. It produces an internally consistent series of meteorological variables including: rainfall, temperature, humidity, wind, sunshine, as well as a derivation of potential evapotranspiration. The publically available version of this generator through the UKCP09 data portal produces base line data and up to 3000 years of future weather for a given time period, representing 100 samples of possible climate change chosen randomly from the probability density function, with data available at a 5 km resolution for the UK.

It is worth mentioning the outcomes related to EARWIG and UKCP09 available through the PROMETHEUS project [59]. This was a multidisciplinary project for the creation of future reference year weather data: TRYs and DSYs based on the CIBSE methodology [17]. The aim was to support physical models to identify the problems new buildings will face as a result of climate change. PROMETHEUS produced weather files in .epw format for 51 locations around the UK for the 2030's, 2050's, and 2080's decades taking into account the high and medium emissions scenarios. As the approach for creating the files was probabilistic (following the nature of the UKCP09 projections), several files were created for each location and decade; the files were then disaggregated by months, ordered by dry bulb temperature, and then assembled by different percentiles of interest. Similarly, the COPSE project, which had a wider remit than PROMETHEUS, also produced a methodology for the creation of future weather files [52]. COPSE aimed to develop robust methodologies for producing weather data files for assessing building designs in future climates, with particular reference to comfort and energy use. COPSE provided DRYs to be used in the design and sizing of heating and air conditioning systems in buildings. Similar to PROMETHEUS, COPSE also produced TRYs and DSYs. While these COPSE years where designed for the three risk levels of the emissions scenarios (low, medium and high), PROMETHEUS produced probabilistic weather files, providing a choice of five probabilities for each future climate.

At the time of writing the UK Government is planning UKCP18, with the aim of improving the robustness of UK climate projections. UKCP18 focuses on upgrading future climate projections over land through a physical understanding of simulated results. The methodology uses new distributions of possible future changes in weather variables, new spatially coherent projections of climate and downscaled simulations of future climate. It is also likely that the spatial resolution will be lower than 5 km [125].

4.2. Morphing time series of weather data

The alternative to using climate projections to prime a weather generator is to adjust (morph) current weather files [16] or even raw time series data taken at weather stations [107]. The starting point of this method is obtaining high-resolution weather data for a specific site. These data are then morphed using projections from either a global or a regional climate model. This process is often used for the analysis of building energy use or assessing resilience under different future climate scenarios [126].

Morphed weather files use historical observations of weather to represent the present-day climate. This produces meteorologically consistent weather files, but ignores some aspects of future climate change, such as the changing frequency of heat waves. The use of a standard baseline historical time series means that the applicability of any future weather files is constrained by baseline data availability. Additionally, if the baseline data to be morphed is already in the form of a TRY, TMY or similar weather file, then there is also the inherent assumption that any climatic change that occurred between the baseline period of the weather file and the baseline period of the climate projections is negligible. The morphed time series are constructed by ‘shifting’ and ‘stretching’ the observed variables using factors produced from the climate change projections. The morphing processes are [107]:

- A ‘shift’ of x_0, is applied by adding the projected change to the absolute monthly mean, Δx_m. This change of the average value is also known as an absolute change.
  $x = x_0 + \Delta x_m$ (1)
  Where x is a future weather variable which changes in the monthly average but with a variance that remains the same. For instance, a ‘shift’ is commonly applied to the baseline of atmospheric pressure.
- A ‘stretch’ of α_m is represented by scaling a future weather variable. This change to the skewness of the data is also known as a fractional change.
  $x = \alpha_m x_0$ (2)

- The associated monthly variance of the variable x also changes: $\mathrm{Var}(x) = \alpha_m \mathrm{Var}(x_0)$. Despite the changes in the monthly variance, the average remains the same. As an example, solar irradiance needs to be 'stretched' rather than 'shifted' so as not to alter the diurnal cycle between day and night.
- Variables can also undergo a combination of 'shift' and 'stretch' transformations.
  $x = x_0 + \Delta x_m + \alpha_m(x - x_0)$ (3)
  In this case we have changes in both the monthly average and monthly variance. Temperature is a variable typically 'shifted' from a baseline in its mean and also 'stretched' in its diurnal range.

Eames et al. [60] conducted a comparison between morphing methods and the UKCP09 weather generator. The study showed that the weather generator is able to statistically produce weather data consistent with historical observations for several variables. However, some issues were present at hourly time-scales with the distribution of both the sunshine hours and the direct and diffuse radiation. The morphing procedure tends to overestimate extreme data as it approaches maximum and minimum temperatures independently from the mean when they should be correlated. This is one of the reasons why weather generator methods are known to be better suited to investigating extreme temperatures. In addition, weather generators offer a consistent, finer spatial resolution than morphing, which is reliant on the number and locations of weather stations [42, 60] (whereas, at least in theory, a weather generator can be run for any location).

4.3. Limitations of future weather

Ideally, current and future weather files would be correlated so that the building response to the weather files can be directly compared. However, the morhing method and using weather generator method have their own limitations in this regard and with other issues which are discussed below.

Weather generators produce time series of weather data and while many time series can be produced that are augmented with projections of climate change, the user receives no information about the climate change anomaly applied. Additionally, while weather generators can produce weather data at a 5 km spatial resolution the climate change projections have a spatial resolution of 25 km (in the case of UKCP09). Hence while much effort goes into producing weather generators that work at higher and higher resolutions to account for topographical features, the creation of future weather is limited by the spatial resolution of the underlying GCM/RCM data. Furthermore, the distributions of individual variables are truncated to limit extremes. As such, it is unlikely that extreme events such as heat waves and storms will be present in short time-series. This means that while the probabilistic outputs of the UKCP09 weather generator are a powerful tool for examining the effects of the range of possible climate change on buildings, they are not ideal when planning for resilience against extreme events.

In addition to how extreme events are handled in climate projections, there is also the issue of limitations in the representation and parameterisation of climate physics and uncertainty associated with future greenhouse gas emissions. These limitations are handled by the Socio-economic Emission Scenarios (now Representative Concentration Pathways) which deal with climate change uncertainty by proposing several likely future climate outlines [127]. Weather generators should address these uncertainties [108] and some climate projections have been designed to take this into account. This is the case with UKCP09, where large Monte Carlo simulations of climate change factors endow the projections with the ability to develop a weather generator and sampling strategy that can be used to gather the effects of climate model uncertainty [128]. This uncertainty is typically achieved via a multi-model ensemble of various GCMs and observations and modelled within a Bayesian framework [129]. In a Bayesian statistical model, all uncertain quantities are modelled as random variables with *a priori* probability distributions produced from previous knowledge of the system. The 'inverse probability' associated with Bayes' theorem allows us to infer unknown quantities, adapt our models, make predictions and learn from data, by combining prior distributions and likelihood into a posterior distributions of parameters.

The use of climate change anomalies to morph historic observations of weather or standard composite weather files also has its own set of limitations in addition to those already mentioned above for the climate change projections. Primary among these is the availability of the baseline weather time series both in terms of spatial resolution and the time period of observations. Climate change anomalies are typically defined according to a baseline period of 1961-1990; hence using a weather file based upon a different time period (or even a shorter period) has the potential to introduce errors into the process, assuming that climate change happens in a uniform, linear manner. The morphing methodology also has the inherent assumption that the weather patterns will not change in the future. The future weather file will contain identical weather patterns to the base weather file albeit with magnitudes of weather variables shifted and stretched by the morphing algorithms. This means that the future weather years will be comparable to the baseline years. Using many years' worth of data for a single location and morphing them might be an interesting approach to cope with the natural variability of climate. However, due to the issues associated with observations of weather, such as missing data, insufficient variables etc. it would be simpler to use a weather generator which includes climate change projections instead. Indeed, through a suitable stochastic weather generator and a combination of climate model realizations, UKCP09 is able to produce time and space consistent future weather files. In general, RCMs with hourly temporal resolution can be used to successfully produce future weather data sets even for the case of extreme conditions [130] through the corresponding climate change projections. However, there are still limitations when using RCMs as their resolution can vary depending on the world region targeted and model(s) employed. Since working with RCMs with multiple variables is computationally intensive, some authors simplify the variables involved in the weather files to just work with the air temperature [130, 131].

5. Including Extremes

Extreme weather events can impact building performance with significant, and potentially lethal, consequences [132]. These events include floods, droughts, or heatwave events, among others. In this review, we focus specifically on extreme events that modify the thermal environment of the occupants, i.e. heat waves and cold snaps. The Fifth Assessment report of IPCC 2014, AR5, on its summary for policy makers states that it is likely[4] that the frequency of heat waves has already increased in large parts of Europe, Asia and Australia [133]. In addition, it is virtually certain that there will be: (i) more frequent hot, and fewer cold, temperature extremes on daily and seasonal scales and (ii) heat waves with higher frequency and duration, than at present [134]. A comprehensive review of indicators of heat waves can be found in [135].

4 The IPCC reports use calibrated uncertainty language, ranging from "exceptionally unlikely" up to "virtually certain" through several intermediate terms such as "unlikely" or "likely", among others. This represents an attempt to express confidence in their assessment of observed data or future climate projections.

5.1. Modelling extreme events

The standard assessment methods on how extreme weather conditions impact the indoor environment are based on thermal comfort [136, 137]. There is a scarcity of studies that cover indoor conditions during severe weather events. An alternative would be to use physiological models to assess risk under extreme events. These models include indices of standard effective temperature, discomfort, physiological strain, and wet bulb global temperature, among others [138], all of which can provide critical information on the effect of the indoor environment on occupant health. Regarding external conditions, several researchers have looked into the prediction and related causes of severe weather events. A key driver of this research is the need to understand the expected return periods and intensities of extreme events for the rest of the 21st century. This literature can be separated into two streams: (1) the modelling of heat waves as a product of the global weather system; thus using a weather model in conjunction with an RCM [134], and (2) modelling severe events using weather data combined with statistical analysis [139].

From the first group are the findings given by the IPCC report, which represent the scientific consensus. Results on heat wave modelling using weather scale simulations can be found in references [140-143]. These proposals work with indices to assess heat wave magnitudes at global and local scale. The frequency and intensity of heat waves is geographically analysed and several future scenarios are considered depending on GCMs, RCMs, and emissions of greenhouse gases. The second group's use of data-driven modelling has not been exploited to the same extent as weather scale simulations. The most influential works from this group are by Kysely [144], Furrer et al. [145], and the recent works of Herrera et al. based on quantile analyses [11, 146]. Through these approaches, the authors are able to characterise and model the frequency and intensity of heat waves and cold snaps based on data from each location. A significant advantage of this approach is that the computational time required is negligible when compared with that needed for weather simulators.

Despite the output of weather models providing useful information on heat waves and their geographical patterns, there is a difficulty in obtaining proper consistent results from simulators. Data-driven models are however more robust and seem to accurately predict the intensity and frequency of extreme events. The downside of the latter approach is that it is highly localised, and large amounts of data are needed to have a realistic prediction of the events.

5.2. Extreme events in buildings

Depending on its characteristics and operation, a building could mitigate or exacerbate exposure to extreme events [147]. Buildings' heating, ventilation and air conditioning systems (HVAC) are designed to cope with a certain range of historical weather conditions to avoid their over-sizing [148]. The coverage is based on annual percentiles of temperature and humidity and typical values include 99.6% and 99% for heating and 1% and 0.4% for cooling. A particular risk level is then chosen according to the end-use of the building (e.g. 99.6%–0.4% for hospitals) [48, 149, 150]. Therefore, there is an explicit exclusion of truly extreme weather conditions that do not occur most years. Although design calculations for energy services in buildings (such as air conditioning systems) include safety factors, these usually account for commissioning issues or limitations, and should not be relied upon a detailed assessment of extreme weather conditions.

The presence of HVAC systems in dwellings is much more case-specific. Houses in temperate climates such as the UK only have conditioning systems for heating, which are used for the mitigation of cold snaps [151]. The current building stock in much of Europe contains vulnerable households with no-air conditioning (as demonstrated in the heat wave of 2003). In these cases, occupants have to rely on night time ventilation to cool down their homes. However, the combination of urban conditions with low wind speeds and heat island effects and situations in which the night time temperature is also high, makes this is ineffectual, leading to fatal consequences [152]. These difficulties for natural ventilation provision could mean that air conditioning systems in buildings become more prevalent to deal with hotter temperatures [153]. In some regions, there is also a trend towards the adoption of reversible heat pumps, which can provide both cooling and heating [154].

The simulation outputs, regarding the building vulnerability to suffer the consequence of a heat wave, are heavily dependent on the quality of the weather file to drive the external weather signal. As demonstrated earlier, current weather files whether built from historical data or from a weather generator are unlikely to contain an adequate representation of extreme events. One possibility would be, as an advancement of the weather generator approach, to use super-synthetic weather files. Super-synthetic weather files are mathematical approximations of meteorological data sets which can be adapted to have several levels of representativeness of the local weather. Methods such as wavelets or Fourier time series decomposition, are basic examples of approaches to further investigate on the creation of super-synthetic weather files. One of the main objectives of these files would be provide data that is computationally efficient to simulate to coerce the representation of extreme events. They can be designed to simulate any number of heat waves in summer and cold spells in winter to check building performance.

More extreme heat waves in developed countries where most households have access to air conditioning systems have led to a different problem. The simultaneous cooling demand caused by adverse weather conditions can overcome grid capacity and cause blackouts [155]. Hence, cooling systems may become unavailable during heat waves. This suggests that the reliability and robustness of systems also need to be studied at a building level. Passive heat-wave-resistant designs will be more robust, whereas buildings dependant on HVAC systems (such as big commercial buildings with poor envelope design, e.g. large glazed areas) may be rendered unusable and need to be evacuated in such events.

6. Conclusions

This paper presents an overview of the methods used to create current and future weather for the analysis of the built environment. Such weather data is generally used to show compliance with policy and regulations, or to examine design alternatives, however there is a growing need to investigate the resilience of building designs, and buildings, to extreme weather events or to climate change.

From this review it is clear that:

- Current weather files are limited by the weather stations where data is collected and their heterogeneous spatial distribution. While you could just use a specific year such as 2003 to model extremes, for a specific location you would need the year to have been recorded and for the event to be significant for that location. For example, the 2003 heat wave in Paris was more extreme for its location than the same period in Edinburgh.
- Despite the ability to create many realisations of synthetic weather cannot reliably generate extremes because weather generators are primed on observed time series.
- Furthermore, we do not have a consistent definition of what a heat wave or a cold snap is, nor how this might change as the climate warms and building occupants adapt. At the root of this is the issue that weather files are always selected as being typical, or atypical, without

respect to the impact they might have on any building. For example, that they are warmer than normal, rather than they might engender overheating in a particular design.

- Future weather files have the same limitations as the current weather files but are also strongly dependant on the uncertainties associated with climate change projections. The most reliable approach currently is to use climate change projections coupled with a weather generator. However, this method still struggles with modelling the full range in variability of future weather and consequently is not best suited to assess severe events.
- The cityscape where the building is located is often not the same as the weather station where the time series is recorded. This ignores aspects such as reduced wind speed and changing wind directions, which will have an impact on the natural ventilation strategy of a building. Furthermore the effects of the urban heat island are frequently also ignored. This can have a dramatic impact on the timing and magnitude of energy use within buildings [156].

Returning to the seven-part list of requirements presented in Section 1, the mappings shown in Table 2 can be made.

Table 2. Mapping of requirements to approaches. ✔ = yes, ✘ = no, ? = possibly.

	Approach				
Requirement	**Observed weather**	**Morphing**	**Weather generator**	**RCM**	**Super Synthetic**
Contains examples of typical conditions	✔	✔	✔	✘	✘
Contains examples of extreme conditions	?	?	?	✘	✔
At the temporal resolution required by simulation packages	✔	✔	✔	✔	✔
At a geographic resolution that matches changes in weather in the country of interest	✔	✔	✔	✔	✔
Likely to well express the urban climate	✘	✘	✘	✘	?
Contain example of possible future climates	✘	✔	✔	✔	✔
Proven track record with industry	✔	✔	✔	✘	✘

From this mapping it is clear that the two areas in need of most urgent attention are:

(i) improving the applicability of current and future building simulation weather files to the urban setting. This will require a greater understanding of the physics associated with phenomena such as the urban heat island and also how these may change over time. In addition, wind models should be adapted to analyse how increased surface roughness associated with cities affects wind speed and directionality.

(ii) Including extremes, particularly ones that might cause morbidity, mortality or heating/cooling system failures.

It is not clear how standard weather files as described above can easily incorporate all requirements as described in the table. As a result the next generation of weather files will need to be redesigned if uniform files types are to be developed for the world which are fit for purpose. As a roadmap the next set of current and future weather files should; (1) be able to cover countries with a high spatial resolution; (2) be able to incorporate any microclimatic effects such as a local urban heat island; (3) can produce typical years, and also provide “extreme” or “unusual” event years ie with events which can be statistically described and can be shown to stress building designs; (4) can do 1, 2 and 3 for future weather; (5) ensures that meaningful inter-decadal comparisons can be made of the evolution in the performance of buildings.

Acknowledgement

This research has been performed in the project COLBE (The Creation of Localized Current and Future Weather for the Built Environment) funded by the EPSRC (EP/M021890/1, EP/M022099/1).

References

1. Klepeis, N.E., et al., *The National Human Activity Pattern Survey (NHAPS): a resource for assessing exposure to environmental pollutants.* Journal of exposure analysis and environmental epidemiology, 2001. **11**(3): p. 231-252.
2. Lader, D., S. Short, and J. Gershuny, *The time use survey, 2005.* How we spend our time. London: Office for National Statistics., 2006. **15**: p. 2009.
3. Fouillet, A., et al., *Excess mortality related to the August 2003 heat wave in France.* International archives of occupational and environmental health, 2006. **80**(1): p. 16-24.
4. Robine, J.-M., et al., *Death toll exceeded 70,000 in Europe during the summer of 2003.* Comptes rendus biologies, 2008. **331**(2): p. 171-178.
5. Charpentier, A., *On the return period of the 2003 heat wave.* Climatic change, 2011. **109**(3-4): p. 245-260.
6. Directive, H.A.T., *Directive 2002/3/EC of the European Parliament an d of the Council of 12 February 2002 relating to ozone in ambient air.* Officia l Journal of the European Communities L, 2002. **67**: p. 14-30.
7. Laustsen, J., *Energy efficiency requirements in building codes, energy efficiency policies for new buildings.* International Energy Agency (IEA), 2008: p. 477-488.
8. Mavrogianni, A., et al., *Building characteristics as determinants of propensity to high indoor summer temperatures in London dwellings.* Building and Environment, 2012. **55**: p. 117-130.
9. De Wilde, P. and D. Coley, *The implications of a changing climate for buildings.* Building and Environment, 2012. **55**: p. 1-7.
10. McMichael, A.J., R.E. Woodruff, and S. Hales, *Climate change and human health: present and future risks.* The Lancet, 2006. **367**(9513): p. 859-869.
11. Herrera, M., et al., *SAX-Quantile Based Multiresolution Approach for Finding Heatwave Events in Summer Temperature Time Series.* AI Communications, 2016. **29**(6): p. 725-732.
12. Burnand, G., *The study of the thermal behaviour of structures by electrical analogy.* British Journal of Applied Physics, 1952. **3**(2): p. 50.
13. Newton, D., R. James, and D. Bartholomew, *Building energy simulation-a user's perspective.* Energy and Buildings, 1988. **10**(3): p. 241-247.
14. Clarke, J.A., *Energy simulation in building design*. 2001: Routledge.
15. Crawley, D.B. *Creating weather files for climate change and urbanization impacts analysis*. in *Building Simulation*. 2007. IBPSA Beijing.
16. Murphy, J., *Predictions of climate change over Europe using statistical and dynamical downscaling techniques.* International Journal of Climatology, 2000. **20**(5): p. 489-501.
17. Mylona, A., *The use of UKCP09 to produce weather files for building simulation.* Building Services Engineering Research and Technology, 2012. **33**(1): p. 51-62.
18. Levermore, G. and J. Parkinson, *Analyses and algorithms for new Test Reference Years and Design Summer Years for the UK.* Building services engineering research and technology, 2006. **27**(4): p. 311-325.
19. Fernández, M., et al., *Generation and evaluation of typical meteorological year datasets for greenhouse and external conditions on the Mediterranean coast.* International journal of biometeorology, 2015. **59**(8): p. 1067-1081.
20. Kershaw, T., M. Eames, and D. Coley, *Comparison of multi-year and reference year building simulations.* Building Services Engineering Research and Technology, 2010. **31**(4): p. 357-369.
21. Finkelstein, J.M. and R.E. Schafer, *Improved goodness-of-fit tests.* Biometrika, 1971. **58**(3): p. 641-645.

22. Yang, L., et al., *Building energy simulation using multi-years and typical meteorological years in different climates.* Energy Conversion and Management, 2008. **49**(1): p. 113-124.
23. NCDC, *Test Reference Year*, in *Tape Reference Manual, TD-9706*. 1976, National Climate Data Center, U.S. Dept. of Commerce: Asheville, U.S.
24. Festa, R. and C.F. Ratto, *Proposal of a numerical procedure to select reference years.* Solar Energy, 1993. **50**(1): p. 9-17.
25. Lund, H. and S. Eidorff, *Selection methods for production of test reference years*, in *Final Report, Report EUR 7306 EN*. 1981, Thermal Insulation Laboratory, Technical University of Denmark: Lyngby, Denmark.
26. Lund, H., *The Design Reference Year user's manual, Thermal Insulation Laboratory.* Technical University of Denmark: Lyngby, 1995.
27. Eames, M.E., A.P. Ramallo-Gonzalez, and M.J. Wood, *An update of the UK's test reference year: The implications of a revised climate on building design.* Building Services Engineering Research and Technology, 2015. **37**(3): p. 316-333.
28. Bilbao, J., et al., *Test reference year generation and evaluation methods in the continental Mediterranean area.* Journal of Applied Meteorology, 2004. **43**(2): p. 390-400.
29. Wilcox, S. and W. Marion, *Users manual for TMY3 data sets*. 2008: National Renewable Energy Laboratory Golden, CO.
30. Lam, J.C., S.C. Hui, and A.L. Chan, *A statistical approach to the development of a typical meteorological year for Hong Kong.* Architectural Science Review, 1996. **39**(4): p. 201-209.
31. Ecevit, A., B. Akinoglu, and B. Aksoy, *Generation of a typical meteorological year using sunshine duration data.* Energy, 2002. **27**(10): p. 947-954.
32. Yang, L., et al., *A new method to develop typical weather years in different climates for building energy use studies.* Energy, 2011. **36**(10): p. 6121-6129.
33. Hall, I.J., et al., *Generation of a typical meteorological year*. 1978, Sandia Labs., Albuquerque, NM (USA).
34. Marion, W. and K. Urban, *User's Manual for TMY2s: Typical Meteorological Years: Derived from the 1961-1990 National Solar Radiation Data Base*. 1995: National Renewable Energy Laboratory.
35. Stoffel, T.L. and M.D. Rymes, *Production of the weather year for energy calculations version 2 (WYEC2) data files/Discussion.* ASHRAE Transactions, 1998. **104**: p. 487.
36. Thevenard, D.J. and A.P. Brunger, *The development of typical weather years for international locations: part I, algorithms.* ASHRAE Transactions, 2002. **108**: p. 376-383.
37. Huang, Y.J., *Development of 3012 IWEC2 weather files for international locations (RP-1477).* ASHRAE Transactions, 2014. **120**: p. 340-355.
38. Huang, J., *ASHRAE Research Project 1477-RP–Development of over 2,500 weather files for international locations.* Final Report. ASHRAE, 2010.
39. Jentsch, M.F., et al., *Limitations of the CIBSE design summer year approach for delivering representative near-extreme summer weather conditions.* Building Services Engineering Research and Technology, 2014. **35**(2): p. 155-169.
40. Yang, L., H. Yan, and J.C. Lam, *Thermal comfort and building energy consumption implications–a review.* Applied Energy, 2014. **115**: p. 164-173.
41. Holmes, M.J. and J.N. Hacker, *Climate change, thermal comfort and energy: Meeting the design challenges of the 21st century.* Energy and Buildings, 2007. **39**(7): p. 802-814.
42. Coley, D., T. Kershaw, and M. Eames, *A comparison of structural and behavioural adaptations to future proofing buildings against higher temperatures.* Building and Environment, 2012. **55**: p. 159-166.
43. Pernigotto, G., et al., *Analysis and improvement of the representativeness of EN ISO 15927-4 reference years for building energy simulation.* Journal of Building Performance Simulation, 2014. **7**(6): p. 391-410.

44. Kalamees, T., et al., *Development of weighting factors for climate variables for selecting the energy reference year according to the EN ISO 15927-4 standard.* Energy and Buildings, 2012. **47**: p. 53-60.
45. CIBSE Guide J, *Weather, solar and illuminance data.* Chartered Institution of Building Services Engineers, 2002.
46. Hacker, J., S. Belcher, and A. White, *Design Summer Years for London (CIBSE TM49).* 2014, Chartered Institution of Building Services Engineers.
47. Eames, M.E., *An update of the UK's design summer years: Probabilistic design summer years for enhanced overheating risk analysis in building design.* Building Services Engineering Research and Technology, 2016: p. 0143624416631131.
48. CIBSE Guide A, *Environmental design.* The Chartered Institution of Building Services Engineers, London, 2015.
49. ASHRAE, A.S., *Standard 55-2013.* Thermal Environmental Conditions for Human Occupancy, 2013.
50. Kyselý, J. and M. Dubrovský, *Simulation of extreme temperature events by a stochastic weather generator: effects of interdiurnal and interannual variability reproduction.* International Journal of Climatology, 2005. **25**(2): p. 251-269.
51. Jentsch, M.F., M.E. Eames, and G.J. Levermore, *Generating near-extreme Summer Reference Years for building performance simulation.* Building Services Engineering Research and Technology, 2015. **36**(6): p. 503-522.
52. Levermore, G.J., et al., *Deriving and using future weather data for building design from UK climate change projections: an overview of the COPSE Project.* 2012.
53. Watkins, R., G. Levermore, and J. Parkinson, *The design reference year–a new approach to testing a building in more extreme weather using UKCP09 projections.* Building Services Engineering Research and Technology, 2013. **34**(2): p. 165-176.
54. Du, H., C. Underwood, and J. Edge, *Generating design reference years from the UKCP09 projections and their application to future air-conditioning loads.* Building Services Engineering Research and Technology, 2012. **33**(1): p. 63-79.
55. Ferrari, D. and T. Lee. *Beyond TMY: climate data for specific applications*. in *Proceedings 3rd International Solar Energy Society conference–Asia Pacific region (ISES-AP-08)*. 2008. Citeseer.
56. Narowski, P., M. Janicki, and D. Heim. *Comparison of Untypical Meteorological Years (UMY) and their influence on building energy performance simulations*. in *Proc. of Conference "Building Simulation–BS2013", Le Bourget-du-Lac*. 2013.
57. Liu, C., et al., *Future probabilistic hot summer years for overheating risk assessments.* Building and Environment, 2016. **105**: p. 56-68.
58. Matzarakis, A. and B. Amelung, *Physiological equivalent temperature as indicator for impacts of climate change on thermal comfort of humans*, in *Seasonal forecasts, climatic change and human health*. 2008, Springer. p. 161-172.
59. Eames, M., T. Kershaw, and D. Coley, *On the creation of future probabilistic design weather years from UKCP09.* Building Services Engineering Research and Technology, 2010. **32**(2): p. 127-142.
60. Eames, M., T. Kershaw, and D. Coley, *A comparison of future weather created from morphed observed weather and created by a weather generator.* Building and Environment, 2012. **56**: p. 252-264.
61. Cai, Y. and D. Hames, *Minimum sample size determination for generalized extreme value distribution.* Communications in Statistics—Simulation and Computation®, 2010. **40**(1): p. 87-98.
62. Fatichi, S., V.Y. Ivanov, and E. Caporali, *Simulation of future climate scenarios with a weather generator.* Advances in Water Resources, 2011. **34**(4): p. 448-467.
63. Furrer, E.M. and R.W. Katz, *Improving the simulation of extreme precipitation events by stochastic weather generators.* Water Resources Research, 2008. **44**(12).

64. Ailliot, P., et al., *Stochastic weather generators: an overview of weather type models.* Journal de la Société Française de Statistique, 2015. **156**(1): p. 101-113.
65. Richardson, C.W. and D.A. Wright, *WGEN: A model for generating daily weather variables*. 1984, US Department of Agriculture, Agricultural Research Service Washington, DC.
66. Wallis, T.W. and J.F. Griffiths, *An assessment of the weather generator (WXGEN) used in the erosion/productivity impact calculator (EPIC).* Agricultural and Forest Meteorology, 1995. **73**(1): p. 115-133.
67. Nicks, A. and G. Gander. *CLIGEN: A weather generator for climate inputs to water resource and other models*. in *Proc. Fifth Int. Conf. on Computers in Agriculture*. 1994.
68. Hayhoe, H. and D. Stewart, *Evaluation of CLIGEN and WXGEN weather data generators under Canadian conditions.* Canadian Water Resources Journal, 1996. **21**(1): p. 53-67.
69. Kou, X., et al., *Validation of the weather generator CLIGEN with daily precipitation data from the Loess Plateau, China.* Journal of hydrology, 2007. **347**(3): p. 347-357.
70. Stöckle, C., et al. *ClimGen: a flexible weather generation program*. in *2nd International Symposium Modelling Cropping Systems. Florence, Italy*. 2001.
71. Castillo, E., et al., *Extreme value and related models with applications in engineering and science*. 2005: Wiley Hoboken, NJ.
72. El Adlouni, S., B. Bobée, and T. Ouarda, *On the tails of extreme event distributions in hydrology.* Journal of Hydrology, 2008. **355**(1): p. 16-33.
73. Dubrovsky, M., *Met&Roll: the stochastic generator of daily weather series for the crop growth model.* Meteorological Bulletin, 1996. **49**: p. 97-105.
74. Jones, P. and M. Salmon, *Development and integration of a stochastic weather generator into a crop growth model for European agriculture.* MARS Project, Final Report to Institute of Remote Sensing Applications, Agricultural Information Systems (ISPRA)., Contract, 1995(5631-93): p. 12.
75. Watts, M., C. Goodess, and P. Jones, *The CRU Daily Weather Generator: BETWIXT technical Briefing Note 1.* The Climate Research Unit, 2004.
76. Chen, J., F.P. Brissette, and R. Leconte, *A daily stochastic weather generator for preserving low-frequency of climate variability.* Journal of hydrology, 2010. **388**(3): p. 480-490.
77. Chen, J., F. Brissette, and R. Leconte, *WeaGETS–a Matlab-based daily scale weather generator for generating precipitation and temperature.* Procedia Environmental Sciences, 2012. **13**: p. 2222-2235.
78. Kreienkamp, F., A. Spekat, and W. Enke, *The weather generator used in the empirical statistical downscaling method, WETTREG.* Atmosphere, 2013. **4**(2): p. 169-197.
79. Yiou, P., *Anawege: a weather generator based on analogues of atmospheric circulation.* Geoscientific Model Development, 2014. **7**(2): p. 531-543.
80. Zorita, E., et al., *Stochastic characterization of regional circulation patterns for climate model diagnosis and estimation of local precipitation.* Journal of Climate, 1995. **8**(5): p. 1023-1042.
81. Yiou, P., et al., *Ensemble reconstruction of the atmospheric column from surface pressure using analogues.* Climate dynamics, 2013. **41**(5-6): p. 1333-1344.
82. Le Bras, J. and V. Masson, *A fast and spatialized urban weather generator for long-term urban studies at the city-scale.* Frontiers in Earth Science, 2015. **3**: p. 27.
83. Remund, J. and S. Kunz, *METEONORM: Global meteorological database for solar energy and applied climatology*. 1997: Meteotest.
84. Remund, J., et al. *The use of Meteonorm weather generator for climate change studies*. in *10th EMS Annual Meeting, 10th European Conference on Applications of Meteorology (ECAM) Abstracts, held Sept. 13-17, 2010 in Zürich, Switzerland. http://meetings.copernicus.org/ems2010/, id. EMS2010-417*. 2010.
85. Ebrahimpour, A. and M. Maerefat, *A method for generation of typical meteorological year.* Energy Conversion and Management, 2010. **51**(3): p. 410-417.

86. Semenov, M.A., E.M. Barrow, and A. Lars-Wg, *A stochastic weather generator for use in climate impact studies.* User Manual, Hertfordshire, UK, 2002.
87. King, L.M., A.I. McLeod, and S.P. Simonovic, *Improved weather generator algorithm for multisite simulation of precipitation and temperature.* JAWRA Journal of the American Water Resources Association, 2015. **51**(5): p. 1305-1320.
88. Soltani, A., N. Latifi, and M. Nasiri, *Evaluation of WGEN for generating long term weather data for crop simulations.* Agricultural and forest meteorology, 2000. **102**(1): p. 1-12.
89. Semenov, M.A., et al., *Comparison of the WGEN and LARS-WG stochastic weather generators for diverse climates.* Climate research, 1998. **10**(2): p. 95-107.
90. Ababaei, B., et al., *Evaluation of a stochastic weather generator in different climates.* Computer and Information Science, 2010. **3**(3): p. 217-229.
91. Smith, S.T. and V.I. Hanby, *Methodologies for the generation of design summer years for building energy simulation using UKCP09 probabilistic climate projections.* Building Services Engineering Research and Technology, 2012. **33**(1): p. 9-17.
92. Goodess, C., et al., *Representing climate and extreme weather events in integrated assessment models: a review of existing methods and options for development.* Integrated Assessment, 2003. **4**(3): p. 145-171.
93. Semenov, M.A., *Simulation of extreme weather events by a stochastic weather generator.* Climate Research, 2008. **35**(3): p. 203-212.
94. Chen, J. and F.P. Brissette, *Comparison of five stochastic weather generators in simulating daily precipitation and temperature for the Loess Plateau of China.* International Journal of Climatology, 2014. **34**(10): p. 3089-3105.
95. Wilby, R., et al., *Evidence needed to manage freshwater ecosystems in a changing climate: turning adaptation principles into practice.* Science of the Total Environment, 2010. **408**(19): p. 4150-4164.
96. Christierson, B.v., J.-P. Vidal, and S.D. Wade, *Using UKCP09 probabilistic climate information for UK water resource planning.* Journal of Hydrology, 2012. **424**: p. 48-67.
97. Ebinger, J.O., *Climate impacts on energy systems: key issues for energy sector adaptation*. 2011: World Bank Publications.
98. Nakicenovic, N., et al., *Special report on emissions scenarios, working group III, Intergovernmental Panel on Climate Change (IPCC).* Cambridge University Press, Cambridge, 595pp. ISBN 0, 2000. **521**(80493): p. 608.
99. Girod, B., et al., *The evolution of the IPCC's emissions scenarios.* Environmental science & policy, 2009. **12**(2): p. 103-118.
100. Pachauri, R.K., et al., *Climate Change 2014: Synthesis Report. Contribution of Working Groups I, II and III to the Fifth Assessment Report of the Intergovernmental Panel on Climate Change*, ed. R.K. Pachauri and L. Meyer. 2014, Geneva, Switzerland: IPCC. 151.
101. Moss, R.H., et al., *The next generation of scenarios for climate change research and assessment.* Nature, 2010. **463**(7282): p. 747-756.
102. Grotch, S.L. and M.C. MacCracken, *The use of general circulation models to predict regional climatic change.* Journal of Climate, 1991. **4**(3): p. 286-303.
103. Nazarenko, L., et al., *Future climate change under RCP emission scenarios with GISS ModelE2.* Journal of Advances in Modeling Earth Systems, 2015. **7**(1): p. 244-267.
104. Van der Linden, P. and J. Mitchell, editors, *ENSEMBLES: Climate Change and its Impacts: Summary of research and results from the ENSEMBLES project.* Met Office Hadley Centre, FitzRoy Road, Exeter EX1 3PB, UK, 2009. **160**.
105. Mearns, L., et al., *Climate change projections of the North American regional climate change assessment program (NARCCAP).* Climatic Change, 2013. **120**(4): p. 965-975.
106. Jacob, D., et al., *EURO-CORDEX: new high-resolution climate change projections for European impact research.* Regional Environmental Change, 2014. **14**(2): p. 563-578.
107. Belcher, S., J. Hacker, and D. Powell, *Constructing design weather data for future climates.* Building Services Engineering Research and Technology, 2005. **26**(1): p. 49-61.

108. Glenis, V., et al., *A transient stochastic weather generator incorporating climate model uncertainty.* Advances in Water Resources, 2015. **85**: p. 14-26.
109. Jentsch, M.F., et al., *Transforming existing weather data for worldwide locations to enable energy and building performance simulation under future climates.* Renewable Energy, 2013. **55**: p. 514-524.
110. Meehl, G., et al., *The WCRP CMIP3 multi-model dataset: A new era in climate change research.* Bulletin of the American Meteorological Society, 2007. **88**: p. 1383-1394.
111. Taylor, K.E., R.J. Stouffer, and G.A. Meehl, *An overview of CMIP5 and the experiment design.* Bulletin of the American Meteorological Society, 2012. **93**(4): p. 485.
112. Meinshausen, M., et al., *The RCP greenhouse gas concentrations and their extensions from 1765 to 2300.* Climatic change, 2011. **109**(1-2): p. 213-241.
113. Nemani, R., et al., *Collaborative supercomputing for global change science.* Eos, Transactions American Geophysical Union, 2011. **92**(13): p. 109-110.
114. Thrasher, B., et al., *Downscaled climate projections suitable for resource management.* Eos, Transactions American Geophysical Union, 2013. **94**(37): p. 321-323.
115. Stoner, A.M., et al., *An asynchronous regional regression model for statistical downscaling of daily climate variables.* International Journal of Climatology, 2013. **33**(11): p. 2473-2494.
116. Hulme, M., J. Turnpenny, and G.J. Jenkins, *Climate change scenarios for the United Kingdom: the UKCIP02 briefing report*. 2002: Tyndall Centre for Climate Change Research, School of Environmental Sciences, University of East Anglia Norwich, UK.
117. Fowler, H. and C. Kilsby. *Future increases in UK water resource drought projected by a regional climate model*. in *Proceedings of the BHS International Conference on Hydrology: Science & Practice for the 21st Century*. 2004.
118. Kilsby, C., et al., *A daily weather generator for use in climate change studies.* Environmental Modelling & Software, 2007. **22**(12): p. 1705-1719.
119. Murphy, J.M., et al., *UK climate projections science report: climate change projections.* 2009.
120. Jenkins, G., et al., *UK Climate Projections science report: Climate change projections*, in *UK09 Briefing Report*. 2009, UK Climate Projections. p. 10.
121. Sexton, D.M., et al., *Multivariate probabilistic projections using imperfect climate models part I: outline of methodology.* Climate Dynamics, 2012. **38**(11-12): p. 2513-2542.
122. Sexton, D.M. and J.M. Murphy, *Multivariate probabilistic projections using imperfect climate models. Part II: robustness of methodological choices and consequences for climate sensitivity.* Climate Dynamics, 2012. **38**(11-12): p. 2543-2558.
123. Betts, R. and M. Best, *Relative impacts of radiative forcing, landscape effects and local heat sources on simulated climate change in urban areas.* Report number UKCIP BETWIXT Project Technical Briefing Note, 2004. **6**.
124. Kilsby, C., A. Moaven-Hashemi, and P. O'Connell. *Simulation of rainfall extremes: fitting to observed annual maxima*. in *First International Conference on Flood Risk, Institute of Mathematics and its Applications, University of Bath, UK*. 2004.
125. UK Met Office, *Technical approach to the delivery of new UK Climate Projections (UKCP18 project)* Technical briefing note, 2016. **V1**.
126. Chan, A., *Developing future hourly weather files for studying the impact of climate change on building energy performance in Hong Kong.* Energy and Buildings, 2011. **43**(10): p. 2860-2868.
127. Lempert, R., et al., *Characterizing climate-change uncertainties for decision-makers. An editorial essay.* Climatic Change, 2004. **65**(1): p. 1-9.
128. Jones, P., et al., *Perturbing a weather generator using change factors derived from regional climate model simulations.* Nonlinear Processes in Geophysics, 2011. **18**(4): p. 503-511.
129. Tebaldi, C., et al., *Quantifying uncertainty in projections of regional climate change: A Bayesian approach to the analysis of multimodel ensembles.* Journal of Climate, 2005. **18**(10): p. 1524-1540.

130. Nik, V.M., *Making energy simulation easier for future climate–Synthesizing typical and extreme weather data sets out of regional climate models (RCMs).* Applied Energy, 2016. **177**: p. 204-226.
131. Cox, R.A., et al., *Simple future weather files for estimating heating and cooling demand.* Building and Environment, 2015. **83**: p. 104-114.
132. Vandentorren, S., et al., *Mortality in 13 French cities during the August 2003 heat wave.* American Journal of Public Health, 2004. **94**(9): p. 1518-1520.
133. IPCC, *Climate Change 2013: Summary for policy makers. In: Climate Change 2013: The Physical Science Basis. constribution of Working Group I*. 2013, Intergovernmental Panel on Climate Change
134. Meehl, G.A. and C. Tebaldi, *More intense, more frequent, and longer lasting heat waves in the 21st century.* Science, 2004. **305**(5686): p. 994-997.
135. Perkins, S.E. and L.V. Alexander, *On the Measurement of Heat Waves.* Journal of Climate, 2013. **26**(13): p. 4500-4517.
136. Bateson, A., *Comparison of CIBSE thermal comfort assessments with SAP overheating assessments and implications for designers.* Building Services Engineering Research and Technology, 2016. **37**(2): p. 243-251.
137. Holmes, S.H., T. Phillips, and A. Wilson, *Overheating and passive habitability: indoor health and heat indices.* Building Research & Information, 2016. **44**(1): p. 1-19.
138. De Freitas, C. and E. Grigorieva, *A comprehensive catalogue and classification of human thermal climate indices.* International journal of biometeorology, 2015. **59**(1): p. 109-120.
139. Branković, Č., et al., *Downscaling of ECMWF ensemble forecasts for cases of severe weather: Ensemble statistics and cluster analysis.* Monthly weather review, 2008. **136**(9): p. 3323-3342.
140. Russo, S., et al., *Magnitude of extreme heat waves in present climate and their projection in a warming world.* Journal of Geophysical Research-Atmospheres, 2014. **119**(22): p. 12500-12512.
141. Koffi, B. and E. Koffi, *Heat waves across Europe by the end of the 21st century: multiregional climate simulations.* Climate Research, 2008. **36**(2): p. 153-168.
142. Fischer, E.M. and C. Schar, *Consistent geographical patterns of changes in high-impact European heatwaves.* Nature Geoscience, 2010. **3**(6): p. 398-403.
143. Easterling, D.R., et al., *Climate extremes: Observations, modeling, and impacts.* Science, 2000. **289**(5487): p. 2068-2074.
144. Kysely, J., *Recent severe heat waves in central Europe: how to view them in a long-term prospect?* International Journal of Climatology, 2010. **30**(1): p. 89-109.
145. Furrer, E.M., et al., *Statistical modeling of hot spells and heat waves.* Climate Research, 2010. **43**(3): p. 191-205.
146. Herrera, M., et al., *Quantile regression ensemble for summer temperatures time series and its impact on built environment studies*, in *8th International Congress on Environmental Modelling and Software, iEMSs*, Sabine Sauvage, Jose-Miguel Sanchez-Perez, and A. Rizzoli, Editors. 2016: Toulouse, France. p. 602-609.
147. Hacker, J., et al., *Climate change and the indoor environment: impacts and adaptation*. 2005: Chartered Institution of Building Services Engineers. 29.
148. ASHRAE, *ASHRAE Handbook: Fundamentals*. 2009, Tullie Circle, N.E.: American Society of Heating, Refrigerating & Air-Conditioning Engineers, Incorporated.
149. UK Dept. of Health, *Heating and ventilation of health sector buildings.* Health technical memoranda, 2007. **03-01**.
150. Mortimer, D., *Controlling humidity for healthier hospitals.* Health estate, 2015. **69**(1): p. 39-43.
151. National-Statistics, *English housing survey 2014 to 2015: headline report*. 2016, Dept. for Communities and Local Gov. UK.

152. Sakka, A., et al., *On the thermal performance of low income housing during heat waves.* Energy and Buildings, 2012. **49**: p. 69-77.
153. Chua, K., et al., *Achieving better energy-efficient air conditioning–a review of technologies and strategies.* Applied Energy, 2013. **104**: p. 87-104.
154. Vellei, M., *A review of factors affecting environmental and economic life-cycle performance for electrically-driven heat pumps*, in *JRC Science and Policy Reports*. 2014, European Commission: Luxembourg. p. 36.
155. Al-Salim, K., I. Andonovic, and C. Michie, *Cyclic blackout mitigation and prevention using semi-dispatchable standby generation and stratified demand dispatch.* Sustainable Energy, Grids and Networks, 2015. **4**: p. 29-42.
156. Futcher, J.A., T. Kershaw, and G. Mills, *Urban form and function as building performance parameters.* Building and environment, 2013. **62**: p. 112-123.

Figure Legends

- Page 3 - Fig. 1. Sample of .epw weather file, showing header information and weather data form line 9.
- Page 10 - Fig. 2. Flowchart of different stages for modelling daily precipitation in a typical weather generator